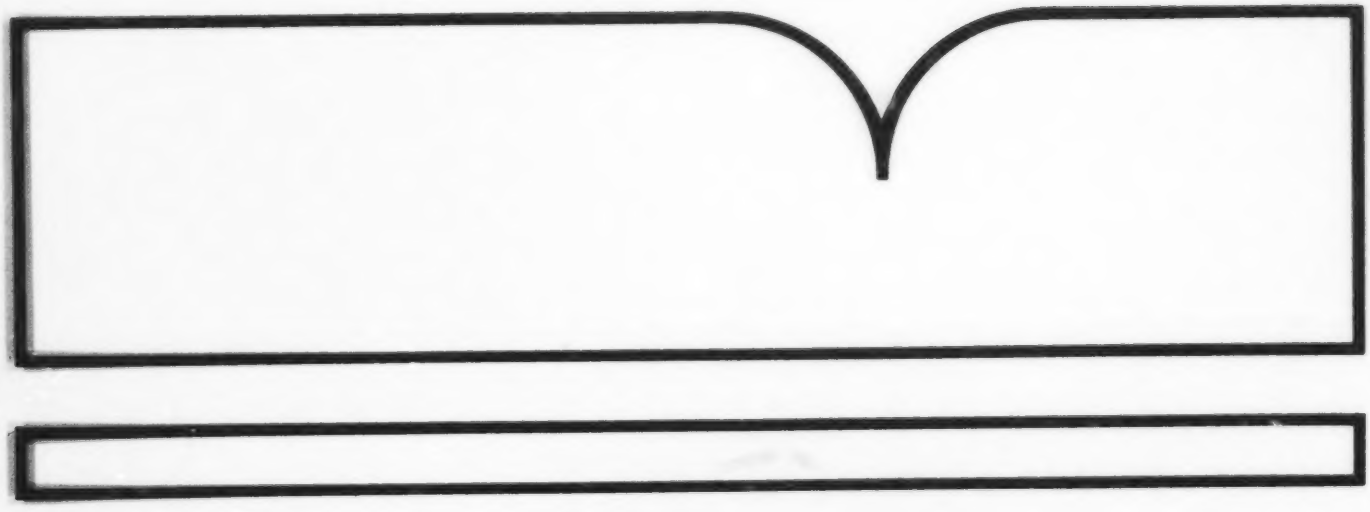
PB87-203089

Analytical Method Comparisons by Estimates of Precision and Lower Detection Limit

(U.S.) Environmental Monitoring Systems Lab.
Research Triangle Park, NC

1986



U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completi

1. REPORT NO.	2.	3. PB87-203089
EPA/600/J-86/377		

4. TITLE AND SUBTITLE	5. REPORT DATE
ANALYTICAL METHOD COMPARISONS BY ESTIMATES OF PRECISION AND LOWER DETECTION LIMIT (Journal Version)	1986
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
David M. Holland and Frank F. McElroy	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
ENVIRONMENTAL MONITORING SYSTEMS LABORATORY OFFICE OF RESEARCH AND DEVELOPMENT U.S. ENVIRONMENTAL PROTECTION AGENCY RESEARCH TRIANGLE PARK, NC 27711	11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Same as Item (9)	14. SPONSORING AGENCY CODE EPA/600/08

15. SUPPLEMENTARY NOTES

REF: Environmental Science & Technology, 20(11):1157-1161, November 1986

16. ABSTRACT

This paper describes the use of principal component analysis to estimate the operating precision of several different analytical instruments or methods simultaneously measuring a common sample of a material whose actual value is unknown. This approach is advantageous when none of the analytical instruments or methods is sufficiently superior to the others to serve as a reference. It also avoids the need for costly experimental designs requiring testing of duplicate instruments. Principal component analysis utilizes the results from tests of three or more analytical instruments or methods measuring common or identical material samples to provide a composite reference value for the material that can be assumed to approximate the true analytical value. From these composite values, estimates of the operating precision for each of the instruments or methods can be developed.

As an example application, this technique is used to estimate the operating precision of six commercially available chemiluminescence analyzers of the type used to measure one-hour average nitrogen dioxide concentrations in the atmosphere. The test group of analyzers included duplicate units of one of the analyzer types, thereby allowing a comparison of the estimate of precision derived from principal component analysis with that obtained from a more conventional technique. The two precision estimates agree closely, corroborating the results of principal component analysis for these data. For each sampler, a measure of the detection limit, defined as the least amount which is detectable with a given degree of certainty, is also provided.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
REPRODUCED BY U.S. DEPARTMENT OF COMMERCE NATIONAL TECHNICAL INFORMATION SERVICE SPRINGFIELD, VA 22161	Principal Component Analysis Lower Detection Limit	

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	6
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

EPA/600/J-86/377
JOURNAL ARTICLE

Published in Environmental Science & Technology, November 1986, pp. 1157-1161, by the American Chemical Society

Analytical Method Comparisons by Estimates of Precision and Lower Detection Limit

David M. Holland[*,†] **and Frank F. McElroy**[‡]

Monitoring and Assessment Division and Quality Assurance Division, Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711

■ Principal component analysis (PCA) can be used to estimate the operating precision of several different analytical instruments or methods simultaneously measuring a common sample whose actual value is unknown. This approach is cost-effective when none of the analytical techniques is sufficiently superior to serve as a reference and obviates the need for experimental designs requiring duplicate instruments. PCA provides composite reference values from sample measurements to approximate the true analytical values. From these composite values, estimates of the operating precision can be obtained. This technique is used to estimate the operating precision of six commercial chemiluminescence analyzers used to measure 1-h average ambient nitrogen dioxide concentrations. Precision obtained alternatively from duplicate data available for one of the analyzer types agreed closely with the PCA estimate. For each analyzer, a measure of a detection limit, defined as the lowest concentration that is detectable with a given degree of certainty, is also provided.

Introduction

Frequently, the need arises to compare the performance of a number of different analytical instruments or methods on the basis of simultaneous measurements of common or identical samples of a material whose true value is unknown. The process of obtaining data is such that, for any given sample, there is only one opportunity for measurement, but it is possible for several instruments to make simultaneous measurements of that sample. The property of the material to be measured is variable over time so that the same instrument cannot make replicate observations. Duplicate measurements may be unobtainable due to high costs of testing duplicate measurement systems.

A number of instruments or methods can be compared under these circumstances for a variety of operational and performance characteristics, one of which would likely be measurement errors caused by various operational factors. Measurement errors are usually described in terms of precision and accuracy, which can also be described by random and systematic errors, respectively. Random analytical error (precision) is a measure of method repeatability, and systematic error (accuracy) is usually described as being composed of additive (fixed) and/or multiplicative (proportional) biases.

As an aid in determining relative operating performance, principal component analysis (PCA), as outlined by Lawton et al. (*1*), may be advantageously used to estimate the operating precisions of several analytical instruments under the conditions described above. This approach involves all portions of the entire measurement process and estimates precision of actual or typical analytical measurements. Confidence intervals may be constructed for each precision estimate to facilitate comparison of the instruments on the basis of the precision associated with each one.

A general discussion of PCA and its use in estimating operating precisions and subsequent determination of detection limits is provided. The procedures are then applied to estimate these performance parameters for six different models of commercial atmospheric nitrogen dioxide analyzers.

Background

When an error-free reference method is available, the precision of p methods simultaneously measuring common

[†] Monitoring and Assessment Division.
[‡] Quality Assurance Division.

or identical samples of a material can be obtained by assuming the sample model

$$y_{ij} = \alpha_j + \beta_j x_i + \epsilon_{ij} \quad (1)$$

where y_{ij} is the ith ($i = 1, ..., n$) observation by the jth method ($j = 1, ..., p$), x_i represents the true value for the ith meaurement, α_j represents the fixed bias, and β_j represents the proportional bias for the jth method. The ϵ_{ij} values are unobservable random measurement errors. The random error ϵ_{ij} is assumed to have a normal distribution with zero mean and standard deviation σ_j, where σ_j is the precision of the jth method.

Grubbs (*2*, *3*) and Thompson (*4*) describe techniques to determine precision by restricting $\beta_j = 1$ for $j = 1$ and 2. Requiring $\beta_j = 1$, however, is often overly restrictive. Barnett (*5*) considered the use of pairwise linear structural relationships between a bias-free reference method and the remaining $p - 1$ test methods. He presented maximum likelihood precision estimates for the test methods but did not consider the problem of negative precision estimates. Russell and Bradley (*6*) used a model slightly different from Grubbs to obtain precision estimates. However, the precision estimates as given by Russell and Bradley may be negative even though the parameters themselves must be nonnegative by definition. Thompson and Moore (*7*) proposed a restricted estimation procedure to obtain positive variance estimates by utilizing the constraint that the estimates be nonnegative. However, for the case of $p \geq 4$, the precision estimates are not unique. Draper and Guttman (*8*) utilized a Bayesian approach to ensure obtaining positive precision estimates for two methods. Carter (*9*) presented restricted maximum likelihood estimates of precision and derived the asymptotic distribution of the estimates when one of the p methods can be accepted as a standard.

For studies where there are no a priori reasons to believe that any one analytical method contains less analytical error than any other method, Theobald and Mallinson (*10*) noted the equivalence of eq 1 to a simple factor analysis model with one common factor, and they used maximum likelihood techniques to estimate the model parameters. Carey et al. (*11*) described the use of PCA to eliminate the necessity of choosing a single reference method. PCA uses the composite of p different analysis methods as a surrogate for the reference method. Lawton et al. (*1*) recognized that PCA could also be used (for $p \geq 3$) to define a total error component of each analytical method. Each error component can be viewed as the effective operating precision for an analytical method and can be used to compare the methods under study.

Principal Component Analysis

Let us assume a linear model relating the measurement of the ith sample by the jth instrument, $[C_j]_i$, to the true concentration, $[C_T]_i$:

$$[C_j]_i = \beta_j[C_T]_i + [E_j]_i \quad (2)$$

$[E_j]_i$ is a total measurement error term composed of random error and error due to the effect of the sensitivity of the instrument to interfering factors. This latter error term is usually referred to as the effect of "interfering substances" (*12*) and is characterized by Lawton et al. (*1*) as random bias. The square root of the total error variance is considered to be the operating precision of the jth method. All of the terms in eq 2 are corrected by their arithmetic averages.

If the range of the measurements is large relative to the total error variance (high signal to noise ratio) and measurement errors are assumed to have a normal distribution with zero mean and constant standard deviation, the elements of the first eigenvector (v_k, $k = 1, ..., p$) of the sample covariance matrix of $[C_j]_i$ will provide an efficient estimate of the normalized β_j. The elements of the first eigenvector can be used to construct the first principal component that is a linear combination of the analytical methods that accounts for the maximum possible proportion of total variation in these methods. Thus, the first principal component is the single best summary of total variance. The least-squares estimate of eq 2 can be written as

$$[\hat{C}_j]_i = \hat{v}_j \sum_{k=1}^{p} [C_k]_i \hat{v}_k \quad (3)$$

If a particular instrument, j, is of special interest or can serve as a reference method, the ratio of v_k/v_j ($j \neq k$) estimates the relative proportional bias between these two analytical instruments. This type of analysis is closely related to estimating the slope coefficient in orthogonal linear regression analysis and has been discussed by Mandel (*13*).

An estimate of the effective operating precision for each instrument is

$$\hat{S}_j^2 = \sum_{i=1}^{n} ([C_j]_i - [\hat{C}_j]_i)^2/(n - 2) \quad (4)$$

The divisor $n - 2$ is used because the residuals have $n - 2$ degrees of freedom after the first eigenvector is estimated from the data. However, the expected values of $\hat{S}_j^2$ are related to the true variances (σ_j^2) by a linear relationship (*1*), assuming $[E_j]_i$ to be independent. Thus, $\hat{S}_j^2$ really estimate linear combinations of the true variances rather than σ_j^2. Proper estimates of σ_j^2 can be obtained by matrix inversion, assuming the number of methods is at least three. A method described by Graybill (*14*) can be used to place confidence limits on σ_j^2 with the χ^2 distribution.

Defining Lower Limit of Detection

The analytical chemistry literature reveals numerous and often conflicting definitions and procedures for estimating a lower limit of detection (LLD) for nonrepeated determinations. It is generally agreed that the formulation of a detection limit should account for the uncontrolled, random fluctuations of the entire analytical process and protect against making misclassification errors (*15–19*).

To describe these random perturbations for blank analyses, the first and second moments (mean and standard deviation) of the distribution of blank analyses must be known or precisely estimated. A blank should be as nearly identical with a field sample as possible, but lacks the substance to be measured. In actual practice, the absence of the substance may be relative, and often, adjustments for nonzero concentrations in the blank are required. With these moments, a critical value can be established above which an instrumental signal (S) is accepted as positive indication of the presence of the substance.

This decision process is subject to two kinds of misclassification errors:

(a) Type I error: declaring a response to be from a population of nonblank values when it is from a population of blank values with probability α ($0 \leq \alpha \leq 1$).

(b) Type II error: declaring a response to be from a population of blank values when it is from a population of nonblank values with probability β ($0 \leq \beta \leq 1$).

Usually α is predetermined by experimenters, assuming a normal distribution of blank observations. However, this assumption may lead to incorrect α probability levels when

distributional information of blank observations is unavailable, which commonly occurs. For these cases, Chebyshev's Inequality (*20*) can be utilized to make a statement of the percent of measurements falling within a specified number of standard deviations of their mean, provided the variance is finite. For example, with Chebyshev's Inequality, at least 89% of the blank signals will occur within three standard deviations of the blank mean. Although Chebyshev's Inequality is conservative, it does provide a bound that is independent of the distribution of blank signals. A limited knowledge of the shape of the distribution of blank signals allows the determination of a sharper boundary. If the distribution is unimodal and has high-order contact with the x axis in the extreme negative and positive tails of the distribution, at least 89% of the blank observations will occur within two standard deviations of the blank mean (*21*). Given sufficient data, one could use the empirical blank distribution to assess parametric distribution forms or use the empirical distribution itself. This would allow tighter bounds than the Chebyshev-type bounds.

Thus, this initial bound or critical value can be calculated as

$$\bar{X}_{\text{blank}} + k_\alpha \sigma_{\text{blank}} \tag{5}$$

where $\bar{X}_{\text{blank}}$ is the mean of blank observations, the values of k_α depends on distributional assumptions, and σ_{blank} is the standard deviation of blank observations. The mean of the blank signals is used to account for any constant bias of the analytical instrument.

The critical value can be incorporated into an expression of the detection limit to protect against the second type of decision error. Theoretically, the detection limit is defined as the smallest true value that will be detected with probability $1 - \beta$, where the observed response represents a nonblank value with probability $1 - \alpha$. Mathematically, the detection limit ($S_{\text{detection}}$) can be expressed as

$$S_{\text{detection}} = \bar{X}_{\text{blank}} + k_\alpha \sigma_{\text{blank}} - k_\beta \sigma_{\text{detection}} \tag{6}$$

where k_β is the abscissa of the standardized normal distribution corresponding to the $1 - \beta$ probability level and $\sigma_{\text{detection}}$ (σ_d) is the standard deviation of possible signals when the true value is equal to the detection limit. The detection limit in concentration units ($C_{\text{detection}}$) is obtained from a response signal by reverse interpolation of the calibration function.

If one can assume that a β error rate of 5% is acceptable and the random errors are normally distributed in the immediate neighborhood of the critical value, then $k_\beta = -1.645$. More complete tabulations of k values can be found in most collections of probability tables. Definitions based on standardized normal probabilities imply that the standard deviation (σ_d) is known, which is never the case. However, if a good estimate of σ_d is available with sample size (n, $n \geq 30$), then definitions based on normal probabilities are reasonable.

From the above, it is clear that for a given instrument the critical value depends only on the choice of α, but the detection limit depends on α and β. The detection limit is also associated with a particular set of procedures for an analytical technique. Further, the same set of instructions may lead to different limits in different laboratories or with different operators.

Thus, the critical value represents the minimum observed value above a true blank or background with confidence $1 - \alpha$, and the detection limit represents the minimum true value that can be detected above background with confidence $1 - \alpha$ and not be declared to be a blank with confidence $1 - \beta$. Assuming normal distributions exist, Figure 1 illustrates these probability statements. Note that the standard derivation of blanks is about half of the level of detection. These levels of variability should not be accepted as universally representative of all types of instruments operating under different site conditions. Comparisons among the estimates of detection can be used to evaluate the relative detection capability of each sampler. The chemical literature often defines the critical value as the detection limit used to decide if a detection occurs. The label, limit of determination, is sometimes used to describe the detection limit as defined here.

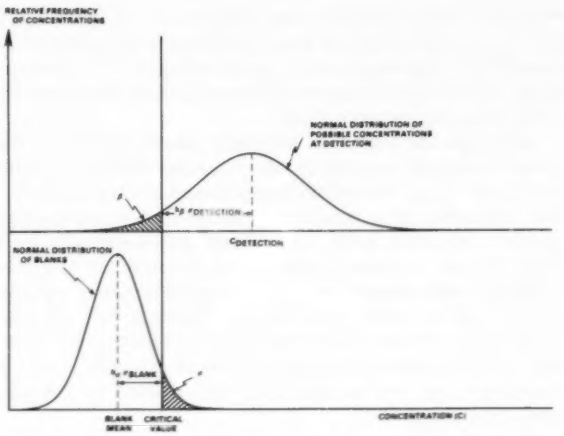


Figure 1. Relationship of α, β, critical value, and lower limit of detection.

Table I. Analyzer Operating Precisions, $\hat{\sigma}_j$, and 95% Confidence Limits

analyzer	operating precision, ppb	lower limit, ppb	upper limit, ppb
Thermo Electron 14 B/E	1.8	1.7	1.9
Bendix 8101C	1.9	1.8	2.0
CSI 1600	2.7	2.6	2.8
Thermo Electron 14 D/E	2.8	2.7	2.9
Beckman 952A	2.9	2.8	3.0
Meloy NA530R	4.2	4.1	4.4

Application to NO_2 Analyzers

As a practical application of the technique, PCA analysis was employed to estimate both the operating precisions and the detection limits of six different models of commercial nitrogen dioxide analyzers (Table I). These analyzers are of the type used by air pollution control agencies to monitor concentrations of NO_2 in the atmosphere, in the range of 0–500 parts per billion (ppb) by volume. Although all of the instruments were based on the chemiluminescence measurement principle for NO_2 (*22*), they differ in design and configuration. Brief descriptions of the analyzers may be found in ref 23. Two identical units of the Bendix Model 8101C were tested to allow comparison of the operating precision estimate derived from the PCA technique with a precision estimate obtained by a more conventional method.

The seven analyzers were tested between April 15 and September 15, 1981, in a typical air monitoring installation located in an urban area of Durham, NC. All of the instruments analyzed identical samples of ambient air drawn from a common, continuously purged manifold, just as they would in an actual air monitoring station. Installation, calibration, operation, and maintenance of the analyzers were in accordance with their operation/instructuion

manuals. The analyzers were individually calibrated once per month, and zero and span checks or adjustments were made 2 or 3 times per week, in accordance with Environmental Protection Agency (EPA) guidelines for operation of air monitoring analyzers.

Although the analyzers normally provide either continuous readings or semicontinuous readings with a cycle time of 1 min or less, individual measurements were defined as 1-h averages, as is common practice in air monitoring. Hourly averages were determined automatically by a system that acquired instantaneous analyzer response readings each minute and calculated the hourly average at the end of each clock hour. These hourly average readings were later converted to concentration measurement on the basis of the calibration data and adjusted, if necessary, for zero or span drift determined by the zero and span check data.

Results and Discussion

It should be emphasized that this analysis deals only with field data, which generally impose a severe restriction on the estimation of bias. Since the true NO_2 concentration in the air at any given time was unknown, it was impossible to determine the absolute accuracy of the methods. The results should be accurate since the field calibrations were performed with National Bureau of Standards traceable materials and standards.

Descriptive summary statistics of the concurrent ambient NO_2 concentration measurements ($n = 1687$) indicate that the central tendency and frequency distribution of the data generated by each analyzer are operationally equivalent. The arithmetic mean values are ~20.0 ppb, standard deviations are ~14.0 ppb, and the data ranged from about 0.0 to 90.0 ppb for each analyzer. This is a result of the generally stable ambient NO_2 average concentrations during the monitoring period. Results from the principal component analysis (appendix) show the high pairwise correlations between the six analyzers. Ratios of the eigenvector elements (for the largest eigenvalue) indicate approximate proportional equivalence between any two analyzers.

Operating Precision Estimates. The Bendix 8101C and Thermo Electron 14/BE had the best estimates (Table I) of effective operating precision (~1.8 ppb). The Beckman 952A, CSI 1600, and Thermo Electron 14/DE operating precisions were roughly equivalent (~2.8 ppb). The Meloy NA530R had the highest level of imprecision (~4.2 ppb). High zero drifts of the Meloy analyzer (23) contributed to the variability of this analyzer. The operating precisions of all the analyzers are well within the EPA specifications (23) of 20.0 ppb at 100.0 ppb and 30.0 ppb at 400.0 ppb. However, in the EPA specifications, concentrations are averaged over the time required to obtain a stable reading (e.g., 5 min). Longer averaging times should result in lower precision estimates due to the stabilizing effect of averaging NO_2 concentrations over 1-h periods. Variability in performance from one instrument to another can be expected because of the complex nature of the analyzers. Even though all the analyzers must be within required specifications, this potential variability must be acknowledged when comparing the precisions of the various models.

The duplicate Bendix measurements can be used to determine an estimate of precision based on a more conventional model. Assuming $\alpha_j = 0$ and $\beta_j = 1$ for $j = 1$ and 2 in eq 1, an estimate of precision can be shown to equal the standard deviation of the differences of the duplicates divided by $\sqrt{2}$ (24). Using this model, the estimate of precision for the Bendix analyzer is 1.92 ppb, which is operationally equivalent to the principal component estimate of 1.87 ppb.

Table II. Lower Detection Limit Estimates

analyzer	blank mean	σ_{blank}	$\sigma_{detection}$	detection limit, ppb[a]
CSI 1600	0.01	0.5	1.7	4
Thermo Electron 14 B/E	−0.02	0.7	2.1	5
Bendix 8101C	0.01	0.4	2.1	4
Beckman 952A	0.04	0.7	2.5	5
Thermo Electron 14 D/E	0.05	0.8	2.4	5
Meloy NA530R	−0.08	3.0	4.0	11

[a] $S_{detection} = k_\alpha \sigma_{blank} - k_\beta \sigma_{detection}$, where $k_\alpha = 1.645$ and $k_\beta = -1.645$ correspond to 5% error rates.

Chemiluminescence NO_2 analyzers are known to be subject to interference from other ambient nitrogen compounds, particularly peroxyacetyl nitrate (PAN) and nitric acid. Although neither of these compounds was measured during the test, they were not likely to be present in significant levels at the test site. Moreover, any minor interference effects during the test would tend to be similar for all collocated analyzers, since they all employed the same type of chemiluminescence measurement principle. The expected minor effect of interferences in this test is apparently confirmed by the close agreement between the two different estimates of precision for the Bendix analyzer model.

Estimation of Detection Limits. In a test such as this one, estimates of the blank mean and variation can be determined from the zero and span data. Frequency histograms of these data indicated a symmetric unimodal distribution. Evaluation of several different goodness-of-fit criteria (25) suggested use of the normal distribution to describe these data. For studies where it is difficult to obtain blank observations, an estimate of blank variation can be approximated from observations near, but above, zero. The blank mean is usually assumed to be zero.

A predetermined region of detection is required to define observations appropriate for determining estimates of σ_d for each analyzer. The limits for this prespecified region were set at

$$\text{critical value} \leq \text{region of detection} \leq 10\ \sigma_i \qquad (7)$$

where σ_i is an estimate of operating precision for the ith analyzer. The basis for the upper limit is described further in ref 26.

Estimates of detection limit were obtained with normal probabilities (Table II). Error rates of 5% for α and β were used ($k_\alpha = 1.645$ and $k_\beta = -1.645$). Excluding the Meloy analyzer, the analyzers could be expected to differentiate true NO_2 concentrations of ~5 ppb from a true blank concentration. The higher detection estimate for the Meloy is due to the higher levels of blank and detection variability. Relaxing the assumption of normally distributed blank observations but maintaining belief in a unimodel, high-order contact distribution allows use of $k_\alpha = 3$ (5% error rate). This larger constant adds ~4 ppb to the Meloy detection estimate and ~0.5–1.0 ppb to the other estimates.

Summary

Principal component analysis of simultaneous measurements by several analytical instruments or methods can be used to obtain estimates of the operating precision for each of the instruments or methods. This type of analysis is advantageous where reference methods are unavailable or where testing of duplicate instruments or methods is not cost efficient. Principal component analysis

uses a composite of measurements from the methods under study as a surrogate for a reference measurement. In addition, an estimate of a lower limit of detection for each instrument or method can be made on the basis of statistical misclassification errors under certain distributional assumptions.

These methods were applied to concurrent 1-hour average measurements generated by six collocated, automated, ambient NO_2 analyzers. Precision estimates varied from ~2 to 4 ppb. A precision estimate obtained from a widely used method based on the differences of measurements generated by two duplicate, collocated analyzers was operationally equivalent to the principal component estimate of operating precision for that analyzer. All of the estimates of detection limit were below ~11 ppb.

Acknowledgments

Vinson L. Thompson of the Quality Assurance Division, Environmental Monitoring Systems Laboratory (Research Triangle Park, NC), served as a project officer of the Equivalency Testing Program, which generated the nitrogen dioxide analyzer performance data analyzed in this paper. We acknowledge the cooperation of Marie Collins, Environmental Monitoring Systems Laboratory (Research Triangle Park, NC), in typing several drafts of the manuscript.

Appendix

Correlation Matrix.

	Bendix 8101C	Thermo Electron 14 B/E	Beckman 952A	CSI 1600	Thermo Electron 14 D/E
Meloy NA530R	0.939	0.935	0.942	0.929	0.930
Bendix 8101C		0.984	0.967	0.974	0.974
Thermo Electron 14 B/E			0.967	0.976	0.974
Beckman 952A				0.964	0.953
CSI 1600					0.958

Sorted Eigenvalues.

1117.8
16.6
9.5
6.8
4.2
2.9

Eigenvector Associated with First (Largest) Eigenvalue.

−0.389
−0.409
−0.403
−0.424
−0.419
−0.402

Registry No. NO_2, 10102-44-0.

Literature Cited

(1) Lawton, W. H.; Sylvestre, E. A.; Young-Ferraro, B. J. *Technometrics* **1979**, *21*, 397–409.
(2) Grubbs, F. E. *J. Am. Stat. Assoc.* **1948**, *43*, 243–264.
(3) Grubbs, F. E. *Technometrics*, **1973**, *15*, 53–66.
(4) Thompson, W. A., Jr. *J. Am. Stat. Assoc.* **1963**, *58*, 474–479.
(5) Barnett, V. D. *Biometrics* **1969**, *25*, 129–142.
(6) Russell, T. S.; Bradley, R. A. *Biometrika*, **1958**, *45*, 111–129.
(7) Thompson, W. A., Jr.; Moore, J. R. *Technometrics* **1963**, *5*, 441–119.
(8) Draper, N.; Guttman, I. *J. Am. Stat. Assoc.* **1975**, *70*, 43–46.
(9) Carter, R. L. *Biometrics* **1981**, *37*, 733–741.
(10) Theobald, C. M.; Mallinson, J. R. *Biometrics* **1978**, *34*, 39–45.
(11) Carey, R. N.; Wold, S.; Westgard, J. O. *Anal. Chem.* **1975**, *47*, 1824–1829.
(12) Mandel, J. *ASTM Stand. News* **1977**, *5*, 17–20.
(13) Mandel, J. *The Statistical Analysis of Experimental Data*; Interscience: New York, 1964.
(14) Graybill, F. A. *An Introduction to Linear Statistical Models*; McGraw-Hill: New York, 1961; Vol. 1.
(15) Altshuler, B.; Pasternack, B. *Health Phys.* **1963**, *9*, 293–298.
(16) Currie, L. A. *Anal. Chem.* **1968**, *40*, 586–593.
(17) Kaiser, H. *Anal. Chem.* **1970**, *42*, 26A–59A.
(18) Oppenheimer, L.; Capizzi, T. P.; Weppelman, R. M.; Mehta, H. *Anal. Chem.* **1983**, *55*, 638–643.
(19) Liteanu, C.; Rîcă, I. *Statistical Theory and Methodology*; Ellis Hardwood: Chichester, U.K., 1980.
(20) Mood, A. M.; Graybill, F. A.; Boes, D. C. *Introduction to the Theory of Statistics*; McGraw-Hill: New York, 1974.
(21) Freeman, H. *Introduction to Statistical Inference*; Addison-Wesley: Reading, MA, 1963.
(22) U.S. Environmental Protection Agency, Title 40, Code of Federal Regulations, Part 50, Appendix F (as amended Dec 1, 1976).
(23) Michie, R. M., Jr.; McElroy, F. F.; Sokash, J. A.; Thompson, V. L.; Fritschel, B. P. "Performance Test Results and Comparative Data for Designated Reference and Equivalent Methods for Nitrogen Dioxide"; U.S. Environmental Protection Agency: Research Triangle Park, NC, 1983; EPA Report No. 600/4-83-019.
(24) Youden, W. J.; Steiner, E. H. *Statistical Manual of the Association of Official Analytical Chemists*; Association of Official Analytical Chemists: Washington, DC, 1975.
(25) Holland, D. M.; Fitz-Simons, T. *Atmos. Environ.* **1982**, *16*, 1071–1076.
(26) "Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry" *Anal. Chem.* **1980**, *52*, 2242–2249.

Received for review November 8, 1985. Revised manuscript received June 13, 1986. Accepted July 3, 1986. This article has not been subjected to Agency review and does not necessarily reflect the views of the Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This article documents research that was conducted under the auspices of the U.S. Environmental Protection Agency (USEPA). As an arm of the government, the USEPA has a paid up, non-exclusive, irrevocable, worldwide license to reproduce, prepare derivative works, and distribute to the public all of the material contained herein, and to grant others acting on its behalf the right to do so. The USEPA grants to the National Technical Information Service, Department of Commerce, the right to act as its agent in the distribution of this material.